

スケジュール疎化による高コスト高報酬行動への 一般化選好の形成¹

田 島 裕 之*

Schedule Thinning Increases Generalized Preference
for High-Cost High-Reward Behavior

Hiroyuki Tajima

高報酬をもたらす強化スケジュールを疎化する手続きによって、反応コストは高いが高報酬を獲得できる行動への一般的な選好を強めることができるかどうかを確認するため、大学生を対象とした実験を行った。参加者はまず、連続強化スケジュールによって低報酬がもたらされる緑色のボタンと間欠強化スケジュールによって高報酬がもたらされる黄色のボタンとの選択を繰り返す課題を行った。各参加者は、間欠強化スケジュールが最初から疎である条件、反応要求の漸増により疎化される条件、間隔要求の漸増により疎化される条件のいずれかに割り当てられた。その後、参加者は全員、連続強化スケジュールによって低報酬がもたらされる赤色のボタンと最初から疎である間欠強化スケジュールによって高報酬がもたらされる青色のボタンの選択を繰り返す課題を行った。前の課題において高報酬をもたらす間欠強化スケジュールが反応要求の漸増によって疎化された人たちの青色ボタンの選択率は、他の人たちよりも高かった。この結果は、報酬が伴わない反応を繰り返すと高報酬を獲得できるという経験が、高コスト高報酬行動への一般的な選好を形成する一因であるということを示唆している。

Key words : response cost, schedule thinning, choice behavior, humans

問題と目的

ある人に大きな良いこと（報酬）をもたらす行動は、その人にとって望ましい行動であると言える。しかしながら、大きな報酬をもたらす行動でも、その報酬獲得までに多数の反応を要する行動（高コスト高報酬行動）は自発しにくい。日常場面では、“（試験に合格するために）勉強する”、“（高額商品を手に入れるために）貯金する”、“（体力をつけるために）運動する”などがこの高コスト高報酬行動に該当すると考えられる。では、本人にとって望ましいにもかかわらず自発しにくい高コスト高報酬行動の自発確率を高めるにはどうすればよいのであろうか。

ある人の高コスト高報酬行動の自発確率を高める方法としてまず考えられるのは、その人に

¹ 本稿は日本心理学会第74回大会で発表したものにデータを追加して再分析を行ったものである。

2014年9月10日受理

* 尚綱学院大学 准教授

行動と報酬との関係を表した言語的記述（ルール）を与えることであろう。しかし、行動に随伴する結果の確率が低いことを示すルールは行動に影響しにくいという指摘があり（Malott, 2005；杉山・島宗・佐藤・マロット・マロット, 1998）、高コスト高報酬行動の行動と報酬との関係を正しく記述するとそのようなルールになってしまう。また、そもそもルールの効果は、ルールを理解できる言語能力とルールに従う一般的傾向とを前提としているため、言語能力の低い人やルールに従う一般的傾向が弱い人に対してはその効果はほとんど期待できない。

高コスト高報酬行動の自発確率を高める他の方法としては、スケジュール疎化（schedule thinning）を上げることができる。スケジュール疎化とは、報酬獲得に要する反応数や時間の長さを徐々に大きくしていく手続きのことである（LeBlanc, Hagopian, Maglieri, & Poling, 2002）。スケジュール疎化は、ある行動の自発確率を、その行動に報酬が伴う頻度が低くなっても（報酬獲得のための反応コストが高くなっても）なるべく維持したい場合に用いられる。スケジュール疎化は言語を必要としないため、人間だけでなく人間以外の動物に対しても適用可能である。

では、スケジュール疎化を用いると行動の反応コストが高くなっても維持されやすいのはなぜであろうか。それは、ある有機体の行動に対してこの手続きを用いると、その有機体はその行動に報酬が間欠的に伴うということを十分に経験するからであると考えられる。ある行動に報酬が間欠強化スケジュールで伴うという経験をさせた有機体と、その行動に報酬が連続強化スケジュールで伴うということを経験させた有機体とで、その行動が消去スケジュール移行後に維持される程度を比較した研究では、前者の方においてその行動が維持されやすいという部分強化消去効果（partial reinforcement extinction effect; PREE）が繰り返し確認されている（例えば、Mowrer & Jones, 1945）。PREE は、間欠強化スケジュールで報酬が伴うという経験をさせた行動と連続強化スケジュールで報酬が伴うという経験をさせた行動との消去スケジュール移行後の選好を調べた研究においても認められている（田島, 2010）。

また、いくつかの研究は、PREE の反応般化という現象、即ち、ある行動に報酬が間欠強化スケジュールで伴うという経験により、その行動だけでなく他の行動も消去スケジュール移行後に維持されやすくなるということを報告している。例えば、Wenrich, Eckman, Moore, & Houston (1967) は、走路走行行動に餌が間欠強化スケジュールで伴う経験をしたラットの方が連続強化スケジュールで伴う経験をしたラットより、餌が連続強化スケジュールで伴っていたバー押し行動が消去スケジュール移行後も維持されやすかったと報告している。同じくラットを用いた McCuller, Wong, & Amsel (1976) の研究では、Wenrich et al. の研究とは反対方向の、バー押し行動から走路走行行動への PREE の般化が認められている。また、Nation, Cooney, & Gartrell (1979) は、大学生を対象とした実験により、PREE がペグを動かす行動からボタン押し行動に般化することを示している。

PREE の有力な説明理論であるフラストレーション理論を提唱した Amsel (1962) は、PREE の反応般化を、反応間の類似性で説明しようとした。この説では、まず、PREE を、(1) ある有機体がある状況下で反応 A に報酬が伴うことを経験すると、レスポナント条件づけの原理でその有機体は状況刺激によって“報酬の予期”が誘発されるようになる、(2) その有機体が“報酬の予期”生起時に、反応 A に報酬が伴わないことを経験すると、それによって一次性的“フラストレーション”が生じ、さらにレスポナント条件づけの原理で、その有機体は状況刺激によって二次性的“フラストレーションの予期”が誘発されるようになる、

(3) その有機体が“フラストレーションの予期”生起時に、反応Aに報酬が伴うことを経験すると、その有機体は“フラストレーションの予期”によって反応Aが喚起されるようになる、という流れで説明する。上記の(2)と(3)は反応Aに報酬が間欠強化スケジュールで伴う場合には生じるが、連続強化スケジュールで伴う場合には生じない。従って、ある有機体が反応Aに報酬が間欠強化スケジュールで伴うことを経験した場合のみ、“フラストレーションの予期”が反応Aを喚起する機能を獲得し、その結果、その有機体の反応Aが消去スケジュール移行後でも維持されやすくなることになる。そして、PREEの反応般化は、Amsel説ではこの“フラストレーションの予期”の行動喚起機能が反応Aに類似した反応に広がったものとして説明される。

Amsel説には、PREEの反応般化があまり似ていない反応間で生じると説明が困難になってしまうという問題点がある。そこで、Wong & Amsel (1976) はAmsel説を拡張し、有機体が“フラストレーションの予期”によって喚起されるようになるものは一般的なフラストレーション対処行動である、とした。彼らは、どのような反応がフラストレーション対処行動に含まれ、どのような反応がそれに含まれないかを明確に示してはいない。ただし、フラストレーション対処行動が、報酬が伴うあらゆる反応を指すのであれば、PREEの反応般化は彼らの説によって説明可能となる。

PREEの反応般化をAmselが行動に先行する刺激の行動喚起機能に着目して説明しようとしたのに対し、Eisenberger (1992) は行動に後続する刺激の行動強化機能に着目して説明しようとした。Eisenbergerは、あらゆる反応には内的な“労力感”が伴い、この“労力感”は反応に要する力が大きいほど、また、反応の生起回数が多くなるほど強くなると考えた。この考えでは、ある行動に報酬が間欠強化スケジュールで伴うと、強い“労力感”に報酬が伴うことになるため、レスポナント条件づけの原理により、強い“労力感”が行動強化機能を獲得する。その結果、強い“労力感”を伴う行動一般が自発しやすくなりPREEの反応般化が生じることになる。

Eisenberger説からは、ある行動に報酬が間欠強化スケジュールで伴うという経験には、PREEの反応般化、即ち、行動一般の消去スケジュール移行後の自発確率を維持する効果だけではなく、最初から報酬が低頻度で伴う高コスト行動一般の自発確率を高める効果があることが示唆される。そこで本研究では、ヒトを対象とし、ある行動に大きな報酬が間欠強化スケジュールで伴うということをスケジュール疎化によって経験させることによって、高コスト高報酬行動一般の自発確率を高めることができるかどうかを調べることを目的とした。高コスト高報酬行動の自発確率は、それが自発していないときに生じる行動を明示的にするため、小さな報酬が連続強化スケジュールで伴う低コスト低報酬行動との選択場面で測定し、そこでの高コスト高報酬行動の選択率を高コスト高報酬行動の自発確率の指標とした。また、高報酬獲得に要する反応数を徐々に増大させるスケジュール疎化と高報酬獲得に要する時間間隔を徐々に増大させるスケジュール疎化とではどちらの方が高コスト高報酬行動一般の自発確率を高めるために効果的であるかも併せて検討した。

方法

実験参加者

A 大学に通う大学生 25 名が実験に参加した。このうち、1 名は後述の訓練セッションにおいて排他的選択を示したため分析対象から除外した。分析対象者 24 名（女性 16 名、男性 8 名）の平均年齢は 19.5 歳（18 歳から 22 歳）であった。

参加者の募集は、A 大学内に“行動実験アルバイト募集”と題したポスターを掲示することにより行った。ポスターには、実験で行う作業の内容、所要時間、謝礼の最低額、実験場所、応募方法が書かれていた。参加希望者には、作業内容、謝礼、所要時間、参加と中断の自由、個人情報の保護、問い合わせ先についての説明が書かれた参加同意書を渡して読んでもらい、参加に同意する場合は、日付、氏名、生年月日、性別、連絡先を記入するように求めた。

実験条件

訓練セッションでの実験課題において、高報酬が伴う行動の強化スケジュールが最初から疎である FR (Fixed Ratio) 9 条件、高報酬獲得に要する反応数が徐々に増大することで強化スケジュールが疎化される PR (Progressive Ratio) 条件、高報酬獲得に要する時間間隔が徐々に増大することによって強化スケジュールが疎化される PI (Progressive Interval) 条件の 3 つを設定した。なお、実験課題は離散試行型であったため、高報酬獲得に要する時間間隔の操作は分や秒ではなく試行数を単位として行った。参加者は実験参加順に基づき 3 名から成る 8 ブロックに分けられ、各ブロック内で参加者を 3 つの条件にランダムに割り付けた。

実験装置

参加者は、室内寸法が幅 1183mm × 奥行 1626mm × 高さ 1964mm の防音室 (YAMAHA 社製 AWA3515H) 内で実験課題を行った。防音室内には、パソコンラックとパイプ椅子があり、パソコンラックには刺激提示兼反応取得用にタッチパネル内蔵 15 型液晶カラーモニター (シャープ社製 LL-151TR) を接続したパーソナルコンピュータ (NEC 社製 PC-MJ30YMZED) を設置した。実験課題の制御は、Microsoft 社の Visual Basic 2005 を用いて作成したプログラムによって行った。

手続き

訓練セッション このセッションでの参加者の課題は、モニター画面に表示された緑色ボタンと黄色ボタン間の選択を繰り返すことであった。

防音室に入室した参加者は、黒字で“スタート”と書かれた白いボタンが画面に表示されているモニターに向かって着席し、次の教示文を読んだ。

実験中、あなたは、画面に表示される正方形のボタンに触れることにより、得点をかせぐことができます。ただし、同時に複数のボタンに触れてはいけません。あなたのかせいだ得点は画面中央に表示されます。実験終了後、あなたのかせいだ得点に応じてお金をお支払いします。1 点は 1 円です。

準備ができましたら、画面の【スタート】ボタンに触れて実験を始めてください。画面

に「これで実験は終了です。」と表示されましたら、実験室から出てきてください。

実験者が防音室のドアを閉めた後、参加者はスタートボタンに触れ、課題を開始した。参加者がスタートボタンに触れるとスタートボタンは消え、その1秒後に第1試行が開始された。各試行は、得点カウンターの左と右に緑色ボタンと黄色ボタンが表示されることによって開始となった。どちらのボタンがどちら側に表示されるかは、毎試行ランダムに決められた。参加者がいずれかのボタンに触れると2つのボタンは消えた。参加者が得点を獲得しなかった場合は、これでその試行が終了した。参加者が得点を獲得した場合は、得点カウンターが0.5×得点秒間白黒反転した状態となり、その間に得点カウンターに表示された数値が0.5秒につき1ずつ増加してその試行が終了した。

全条件において緑色ボタンは1回選択するたびに1点獲得できるようになっていた。黄色ボタンは、FR9条件では9回選択するたびに18点獲得できるようになっていた。PR条件では、黄色ボタンは x 回選択するたびに18点獲得でき、 x の値は1から9まで得点獲得のたびに1ずつ増加するようになっていた。PI条件では、黄色ボタンは $x - 1$ 試行の間隔を空けてから1回選択すると18点獲得することができ、 x の値は1から9まで得点獲得のたびに1ずつ増加するようになっていた。

課題の試行数は270であり、試行間間隔は1秒であった。課題終了後、参加者は別室で獲得した得点に応じた現金を受け取った。

テストセッション このセッションでの参加者の課題は、モニター画面に表示された赤色ボタンと青色ボタン間の選択を繰り返すことであった。全条件において、赤色ボタンは1回選択するたびに1点獲得できるようになっており、青色ボタンは9回選択するたびに18点獲得できるようになっていた。それ以外については、訓練セッションと同じであった。

結果

図1は、訓練セッションにおける各参加者の黄色ボタン選択の強化率を示している。FR9群の黄色ボタンの強化率 ($M = 0.03$) はPR群 ($M = 0.14$)、PI群 ($M = 0.43$) と比べて低くなっているが、これは条件設定による必然的な結果である。なぜなら、FR9条件での黄色ボタン選択の強化率の最大値は、PR条件、PI条件における黄色ボタン選択の強化率の最小値より小さくなるからである。FR9条件における黄色ボタン選択の強化率の最大値は、黄色ボタン選択回数9の倍数となったときの0.11である。これに対して、PR条件における黄色ボタン選択の強化率の最小値は、黄色ボタン選択回数が269のときの0.12である。また、PI条件における黄色ボタンの強化率の最小値は、黄色ボタンを269回連続して選択してから緑色ボタンを1回選択したときの0.12である。そこで、PR群とPI群との比較に対してのみ統計的検定を行った。黄色ボタン選択の強化率に対する群の効果量 r を検定統計量とするランダム化検定の結果、PI群の黄色ボタン選択の強化率はPR群より有意に大きかった ($r = .83$, $N = 16$, $p = .008$)。なお、FR9群8名中5名は黄色ボタン選択の強化率が0であり、黄色ボタン選択に高報酬が間欠強化スケジュールで伴うという経験をまったくしていなかった。

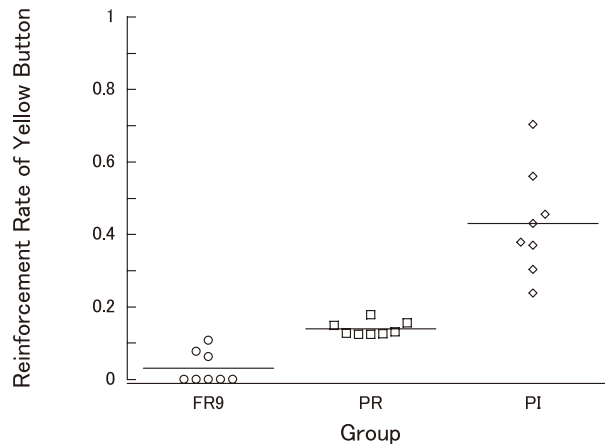


図 1 訓練セッションにおける黄色ボタン選択の強化率。横線は各群の平均値を表す。

図 2 は、訓練セッションにおける各参加者の黄色ボタン選択率を示している。黄色ボタン選択率に対する群の効果量 r を検定統計量としたランダムイゼーション検定 (p 値は Shaffer の法により修正) の結果、PR 群の黄色ボタン選択率 ($M = 63\%$) は FR9 群 ($M = 8\%$) より 5% 水準で有意に大きく ($r = .79, N = 16, p = .023$)、また PI 群 ($M = 26\%$) よりも有意に大きかった ($r = .66, N = 16, p = .023$)。さらに、PI 群の黄色ボタン選択率は FR9 群より有意に大きかった ($r = .61, N = 16, p = .031$)。

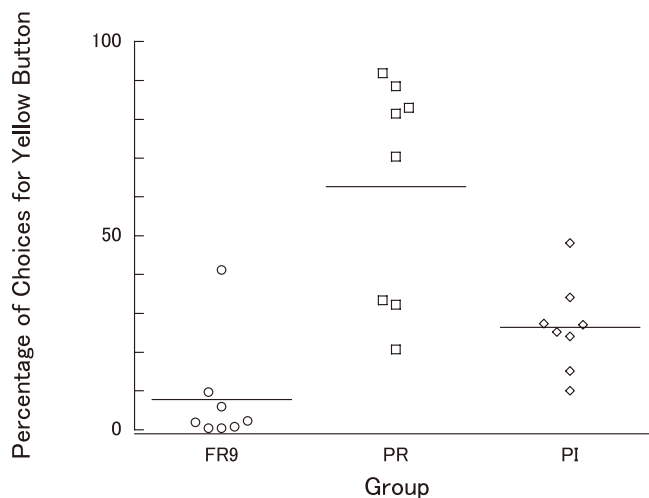


図 2 訓練セッションにおける黄色ボタン選択率。横線は各群の平均値を表す。

PR 条件では黄色ボタン選択回数数が 36 に達すると、それ以降は FR9 条件と同じく、得点に要する黄色ボタン選択反応数が 9 に固定となる。PR 群は全員、得点に要する黄色ボタン選択反応数が訓練セッション内で 9 に達した。そこで、PR 群について、得点に要する黄色ボタン選択反応数が 9 のときの各参加者の黄色ボタン選択率を求め、FR9 群の黄色ボタン選択率と比較した (図 3)。黄色ボタン選択率に対する群の効果量 r を検定統計量とするランダムイ

考察

本研究では、ある行動の報酬獲得に要する反応数を徐々に増加させるという方法には、その行動の自発確率を維持する効果だけでなく、新しい行動選択場面における高コスト高報酬行動の自発確率を高める効果もあるということが確認された。このことは、この方法には高コスト高報酬行動一般の自発確率を高める効果があるということを示唆している。これに対して、ある行動の報酬獲得に要する時間間隔を徐々に増大させるという方法には、そのような効果は認められなかった。そもそも、ある行動の報酬獲得に要する時間間隔を徐々に増加させるという方法は、その行動を維持する効果が弱いようである。PI 群の黄色ボタン選択の強化率の方が PR 群より高かったにもかかわらず、黄色ボタン選択率は PR 群の方が高かったという結果は、行動の自発頻度とその行動に報酬が伴う頻度との相関関係が強いほど、その行動は自発しやすくなるとする反応-強化子相関理論 (Baum, 1973; Green, Kagel, & Battalio, 1987) の主張と一致する。

反応数に基づくスケジュール疎化が新しい選択場面における高コスト高報酬行動の自発確率を高めるという本研究の結果は、PREE の反応般化を反応間の類似性によって説明しようとする Amsel 説では解釈困難である。なぜなら、今回の実験で用いられた反応はすべて、モニター画面上のボタンに触れるという同じ特性のものであったため、Amsel 説はテストセッションにおける特定の選択肢への選好を予測しないからである。また、反応特性にボタンの色を含めたとしても、訓練セッションで黄色ボタンの方をより多く選んでいた PR 群がテストセッションでより多く選んだのは黄色と類似性の高い緑色のボタンの方ではなく、類似性の低い青色のボタンの方であり、Amsel 説によって予測される結果とは反対である。

また、Wong & Amsel 説も本研究の結果を予測することができない。“フラストレーションの予期”によって喚起されるものが、報酬が伴うすべての反応であるとしたら、それには、本研究のテストセッションにおける青色ボタンの選択も赤色ボタンの選択も該当してしまうからである。

さらに、本研究は Eisenberger 説から着想を得て行ったものであるが、それでも本研究の結果は Eisenberger 説では解釈が困難である。今回の実験で用いられた反応はどれもほぼ同じ力を要するものであったため、どの反応もそれに伴う“労力感”はほぼ等しいと考えられる。よって、Eisenberger 説はテストセッションでの特定の選択肢への選好を予測しない。

本研究の結果は、Eisenberger 説を若干修正し、ある行動に報酬が間欠強化スケジュールで伴うという経験によって行動強化機能を獲得したのは“労力感”という内的刺激ではなく無報酬という外的刺激であると考えれば解釈可能である。訓練セッションにおいて、PR 群は 8 名全員が、黄色ボタンの選択に高報酬が間欠強化スケジュールで伴うということ、即ち、無報酬に高報酬が伴うということを経験していた。これに対し、FR9 群ではその経験を 8 名中 3 名しかしておらず、さらにその 3 名のその経験回数は PR 群と比べると少なかった。よって、訓練セッションにおいて無報酬が獲得した行動強化機能の強さは PR 群の方が強く、その結果、テストセッションにおける青色ボタン選択率（高コスト高報酬行動選択率）は PR 群の方が大きくなったと考えられる。本研究ではテストセッションにおいて高コスト高報酬行動の選択率を測定したが、この解釈が妥当かどうかは、テストセッションにおいて無報酬行動の選択率を測定することによって、より直接的に確かめることができるであろう。

また、本研究の結果は、テストセッションにおいてPR群が示した高コスト高報酬行動への選好は、訓練セッションで生成された自己ルールに制御されたルール支配行動（Skinner, 1969）であるという解釈も可能である。ただし、本研究のような結果が得られるためには、そのルールは“黄色いボタンを選択し続けるとやがて高得点が得られる”といった、反応の特性もしくは反応に先行しそれを特定する刺激の特性とその反応の将来の結果との随伴性を記述したものではなく、“得点が得られないボタンの方を選択し続けるとやがて高得点を得られる”といった、反応直後に生じる刺激の特性とその反応の将来の結果との随伴性を記述したものでなければならない。さらに、ある行動がこのようなルールによって制御されるということは、その行動はその直後の刺激に影響されるようになるということを意味するが、このようなルールの機能を行動直後の刺激による行動強化機能と区別することが可能なのかという問題が残る。

文献

- Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discriminative learning: Some recent history and a theoretical extension. *Psychological Review*, **69**, 306-328.
- Baum, W. M. (1973) The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, **20**, 137-153.
- Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, **99**, 248-267.
- Eisenberger, R., & Adornetto, M. (2005). Generalized self-control of delay and effort. *Journal of Personality and Social Psychology*, **51**, 1020-1031.
- Green, L., Kagel, J. H., & Battalio, R. C. (1987) Consumption-leisure tradeoffs in pigeons: Effect of changing marginal rates by varying amount of reinforcement. *Journal of the Experimental Analysis of Behavior*, **47**, 17-28.
- LeBlanc, L. A., Hagopian, L. P., Maglieri, K. A., & Poling, A. (2002) Decreasing the intensity of reinforcement-based interventions for reducing behavior: Conceptual issues and a proposed model for clinical practice. *The Behavior Analysis Today*, **3**, 289-300.
- Malott, R. W. (1986). Self-management. In M. G. Hersen & Rosqvist (Eds.), *Encyclopedia of Behavior Modification and Cognitive Behavior Therapy*. Vol. 1. Newbury Park CA: SAGE Publications. pp. 516-521.
- Mowrer, O. H., & Jones, H. (1945) Habit strength as a function of the pattern of reinforcement, *Journal of Experimental Psychology*, **35**, 293-311.
- McCuller, T., Wong, P. T. P., & Amsel, A. (1976). Transfer of persistence from fixed-ratio barpress training to runway extinction. *Animal Learning and Behavior*, **4**, 53-57.
- Nation, J. R., Cooney, J. B., & Gartrell, K. E. (1979). Durability and generalizability of persistence training. *Journal of Abnormal Psychology*, **88**, 121-136.
- Skinner, B. F. (1969) *Contingencies of Reinforcement: A Theoretical Analysis*. New York: Appleton-Century-Crofts.
- 島宗理, 佐藤方哉, リチャード・W. マロット, アリア・E・マロット (1998). 行動分析学入門産業図書
- 田島裕之. (2010). 消去時の行動持続性: 小強化子による連続強化と大強化子による部分強化の効果の比較, 尚絅学院大学紀要, **60**, 203-207.
- Wenrich, W. W., Eckman, G. E., Moore, M. T., & Houston, D. F. (1967). A transresponse effect of partial reinforcement. *Psychonomic Science*, **9**, 247-248.
- Wong, P. T. P., & Amsel, A. (1976). Prior fixed ratio training and durable persistence in rats. *Animal Learning and Behavior*, **4**, 461-466.